

Subhrato Som

Philadelphia, Pennsylvania 19104

📞 215-240-9510 ✉ ss5654@drexel.edu [in linkedin.com/in/subhratosom/](https://www.linkedin.com/in/subhratosom/) github.com/Subhrato20

Education

Drexel University

Master's of Science in Computer Science

Sep 2023 - May 2025

Philadelphia, PA, USA

Vellore Institute of Technology

Bachelor of Technology in Computer Science and Engineering

Jul 2019 - Jun 2023

Chennai, TN, India

Coursework: Machine Learning, Deep Learning, Data Structures & Algorithms, Software Design, Test Driven Development, Database Management, DevOps, CI/CD, Distributed Systems, Agile Methodology

Technical Skills

Cloud & DevOps: AWS EC2/S3/SageMaker, Docker, Jenkins, Kubernetes

Languages: Python, Java, JavaScript, SQL, C++

Data Engineering: ETL Pipelines, PostgreSQL, MongoDB, Snowflake, Data Modeling

Machine Learning: Scikit-learn, pandas, TensorFlow, PyTorch

Web Development: Flask, Node.js, React, HTML/CSS

System Design: Distributed Systems, Object-Oriented Design, REST APIs

Experience

AI Intern

Jun 2024 - Dec 2024

Mathtech, Inc

Hamilton, NJ

- Built end-to-end data pipeline for RAG system using AWS SageMaker & EC2, implementing document storage, embedding generation, and vector search workflows
- Optimized chunk parsing through dynamic size/overlap adjustment and BERT-based ranking, improving retrieval relevance by 32%
- Designed AWS Cognito integration for secure user authentication and access management
- Implemented batch processing for large-scale document datasets using Python and Milvus vector database

Research Assistant

Feb 2024 - Present

Social-NLP Lab, CCI Drexel University

Philadelphia, PA

- Developed Python-based batch API methods for GPT inference, processing 50k+ Reddit posts for substance abuse pattern analysis
- Performed comparative analysis of LLM performance (LLaMA vs GPT) using statistical methods and human evaluation metrics
- Engineered data preprocessing pipelines for multi-modal meme datasets combining textual and visual features

Projects

Disaster Risk Analytics Platform | Philly Codefest 2024, Chubb Challenge Winner

Apr 2024

- Architected AWS-based data pipeline (EC2/S3) integrating Poisson distribution models with deep learning for financial risk prediction
- Developed Flask API backend serving real-time risk assessments to React frontend with interactive US map visualization
- Implemented automated data ingestion system processing NOAA weather data and historical disaster records

REppeaR (Blockchain Real Estate Platform) | Vellore Institute of Technology

Mar 2022 - Jun 2022

- Designed MongoDB database schema with multi-user access controls for property transaction records
- Built JavaScript blockchain network implementing cryptographic hashing and digital signature verification
- Created ETL pipeline for property data integration from government land records using Python

Question Answering System using BERT | Vellore Institute of Technology

Feb 2022 - Apr 2022

- Engineered Python data preprocessing pipeline for CoQA dataset using NLTK and spaCy libraries
- Fine-tuned BERT model achieving 84.1% accuracy through hyperparameter optimization
- Implemented cosine similarity scoring system for answer relevance evaluation

Activities

Chubb Challenge Winner, Philly Codefest 2024	Apr. 2024
React Summit, GitNataion	June 2022
Open Source Contributor, Girlsript Foundation	Sep. 2021 – Dec 2021
Student Expert, Postman	Aug. 2021 – Jul 2023

Certifications

PostgreSQL for Everybody, University of Michigan	Feb. 2024
Google Data Analytics Professional Certification, Google (Coursera)	Jan. 2022
Microsoft Azure Fundamentals (AZ-900), Microsoft	May. 2022
Data Visualization with Tableau, UC Davis (Coursera)	Feb. 2023